DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES

2007-05

McMASTER UNIVERSITY

Department of Economics
Kenneth Taylor Hall 426
1280 Main Street West
Hamilton, Ontario, Canada
L8S 4M4

http://www.mcmaster.ca/economics/

Ex Post Welfare under Alternative Health Care Systems

John Leach*
McMaster University

May 29, 2007

Abstract

This paper studies the implications of a societal aversion to inequality for the optimal design of a public health care system. Inequality aversion is introduced by postulating a strictly concave ex post social welfare function. Illnesses are characterized by three factors: the agent's health with treatment, the agent's health without treatment, and the cost of treatment. It is shown that the optimal public health care system allocates health care differently than would private health insurance; specifically, people who are relatively unhealthy with and without treatment receive more health care, and people who are relatively healthy with and without treatment receive less health care. The aggregate quantity of health care under the optimal public health care system might be either greater or less than under private health care insurance. If the public health care system is optimally designed, allowing agents to purchase supplementary private health care insurance cannot raise social welfare and is likely to decrease it.

1. Introduction

Although public health care developed largely in response to concerns over equity, these concerns are missing from the existing studies of the allocative effects of health care programs. The intent of the current paper is to correct this omission,

*I should like to thank Tom Crossley, Seungjin Han, Jerry Hurley and Robert Nuscheler for valuable conversations.

with an emphasis on two issues. First, how does inequality aversion affect the structure of a public health care program? Second, does inequality aversion cause societies to prefer public health care to private health care?

There are several ways in which inequality aversion can be modelled. One is Tobin's [11] specific egalitarianism, under which society's preference for relatively equal distributions extends only to particular goods, health care being one example. Another is Pauly's [7] health care externality, under which each agent cares about the other agents' health as well as his own. The externality approach is interesting because it implements specific egalitarianism without forsaking the tools of welfare economics, but it can be analytically cumbersome. A third approach—the one that is adopted here—assumes that each agent cares only about his own health, and that the policy-maker maximizes a social welfare function that is increasing and strictly concave in the individual utilities. The concavity of the social welfare function implies inequality aversion.

Concavity has one other important implication. Agents are uncertain about their own future health, and there are competing concepts of social welfare whenever uncertainty is present. This issue has been addressed by Starr [10], Mirrlees [6], Harris [5] and Hammond [4]. Hammond's discussion is particularly well-suited to the case at hand. He defines the *ex ante social welfare function* as a function W_A whose arguments are the agents' expected utilities, and the *ex post social welfare function* as the expected value of some function W_P whose arguments are the agents' realized utilities. If W_A is a weighted sum of the expected utilities, it is mathematically equivalent to an ex post social welfare function in which W_P is the same weighted sum of the realized utilities. However, the mathematical equivalence between the two functions is lost when W_A (or W_P) is strictly concave, so that we are forced to choose between these concepts.

There does not seem to be a compelling reason for choosing one criterion over the other. Both functions are consistent with non-paternalism. Agents confronted with an uncertain future are forced to maximize their expected utilities, but they do so knowing that only one outcome will be realized, and that their utility will ultimately depend only on that outcome. The issue is simply whether we wish to measure the welfare of the agents before or after the realization. Although Starr opts for the latter approach, arguing that ex ante welfare is a "normative dead end," other investigators have been more circumspect.

Ex post welfare seems to be an appealing criterion in the present case, when there is both uncertainty over future health and aversion to inequality. Should the health care system be designed to allay the agents' fears of future illness, or to ameliorate the effects of illness once it has occurred? Does inequality aversion mean that society is averse to unequal prospects or to unequal outcomes? The first option in each question concerns ex ante welfare, while the second concerns ex post welfare. I favour the second options, and I suspect that I am not alone.

Studies in welfare economics normally assume utilitarian social welfare functions, so that it is not necessary to choose between the two concepts. The redistributive role of health care, for example, has been carefully analyzed by Boadway, Leite-Monteiro, Marchand and Pestieau [1, 2] and Petretto [8].[1] In each case the social welfare function is the sum of expected utilities, or equivalently, the expectation of the sum of the realized utilities. The distinction between ex ante and ex post welfare is not relevant to them.

It will be assumed here that the government maximizes a strictly concave ex post social welfare function. Agents will be assumed to be ex ante identical but uncertain about the future state of health. Health is represented by a triplet whose elements are the state of health without treatment, the state of health with treatment, and the cost of treatment. A production possibility frontier describes the combinations of aggregate health care and aggregate consumption that the economy can produce.

Each health care system is characterized by the kinds of agents that are treated under that system and the kinds that are not treated. The allocation of resources under private health care insurance differs from that under public health care; specifically, public health care shifts health care away from those who are relatively healthy with and without treatment, and towards those who are relatively unhealthy with and without treatment. Aggregate health care could be either greater or smaller under public health care than under private health care. Public health care is preferred to private health care not because it corrects some market failure, but because it has an entirely different objective. Private insurance markets forces individuals to maximize expected (or ex ante) utility while the government, which need not act until every agent's health status is known, is able to allocate resources to maximize an ex post welfare function.

There is a laissez-faire argument that a system of private health insurance, operated in conjunction with the public health care system, must raise welfare: those who purchase private health insurance are better off, while those who do not purchase it are unaffected. This argument does not hold here. A parallel system of private health insurance moves the economy towards the allocation that would occur under a purely private health insurance system, and that system does not

[1] Their work builds on earlier contributions by Rochet [9] and Cremer and Petieau [3].

maximize ex post welfare. An attempt to supplement public health care with private health insurance can have only two outcomes: either no one purchases it, or ex post welfare falls. The latter outcome can arise because the government is forced to take the reaction of the private health care system into account when it designs its own system. This reaction constitutes an additional constraint on the government's maximization problem. That constraint can be binding, and as always, the addition of a binding constraint reduces the maximized value.

The government's first-best policy—its policy in the absence of active private insurers—takes a very simple form: the social value of another unit of health care is equated to the social value of the consumption forgone when this unit is produced. This policy is also its second-best policy—its policy in the presence of active private insurers—if the marginal cost of health care is constant. However, if the marginal cost of health care is increasing, the second-best policy pushes the social value of another unit of health care *below* the social value of the forgone consumption. This policy reduces the amount of health care provided by the private insurers, increasing the government's control over health care resources.

2. The Model

An agent is identified by his type a, where

$$a \equiv (h_0, h_1, m)$$

Here, h_0 is the agent's health without medical treatment, h_1 is health with treatment, and m is the cost of treatment (measured in units of health care). Each element of the sample space of agent types, A, satisfies the conditions

$$0 \leq h_0 < h_1$$

$$0 < m$$

The sample space is assumed to be compact and connected. The utility of each agent is

$$U = h + u(c)$$

where h is health (either h_0 or h_1) and c is consumption. The function u is increasing and concave.

Let $F(X)$ be the measure of the agents whose types are contained in X (where X is any subset of A), and assume that F is differentiable. The set of all types that are not treated is A_0 and the set of all types that are treated is A_1. By definition,

$$F(A_0 \cap A_1) = 0$$
$$F(A_0 \cup A_1) = 1 \tag{1}$$

Health care and consumption goods are produced under constant returns to scale using some vector of inputs. There is a fixed supply of each input. The efficient allocation of these inputs between the industries gives rise to a production possibility frontier that is concave but not necessarily strictly concave. The frontier is

$$C + G(M) = 0 \qquad G' > 0, G'' \geq 0$$

Here, C and M are aggregate consumption and aggregate health care respectively:

$$C \equiv \int_A c dF$$
$$M \equiv \int_{A_1} m dF \tag{2}$$

Aggregate health care is bounded above by

$$\overline{M} \equiv \int_A m dF$$

It is assumed that $G(\overline{M})$ is negative, so that it is possible to treat all of the agents.

The inputs are owned by the agents, who inelastically offer them for sale. Both goods are produced by competitive and privately owned firms. Since the firms are price-takers in both the input and output markets, the total income of the agents is equal to the market value of the produced goods. Their total income (measured in consumption goods) is

$$Y \equiv C + G'(M)M$$

Each agent has an equal claim on this income.

3. Private Health Care Insurance

Assume that the agents, prior to learning their own types, are able to contract with competitive health care insurers. Each insurance policy specifies a set of types for which treatment will be provided. The insurer collects premiums from all insurees, and provides treatment by purchasing health care from the health care producers. Competition among the insurers ensures that the policies will have the following properties:

- No policy will provide health care to a positive measure of types for which

$$h_1 - h_0 < u'(c)G'(M)m$$

  The cost of treating these types would exceed the additional premiums that could be collected by including their treatment in a policy.

- The set of types for which

$$h_1 - h_0 > u'(c)G'(M)m \tag{3}$$

  and for which no policy provides coverage is measure zero. If a positive measure of such types existed, some insurer would be able to earn profits by offering a policy that covered them.

- Competition among insurers reduces each insurer's profits to zero. That is, the total premiums collected by an insurer are equal to the cost of providing treatment to its insurees.

The last observation implies that, in equilibrium, every agent benefits from the purchase of insurance if the set of types satisfying (3) has positive measure. The total premiums collected by the insurers are equal to $G'(M)M$. After paying these premiums, the agents have just enough income to purchase all of the consumption goods. Since each agent has equal income,

$$c = C = -G(M) \tag{4}$$

for all agents.

Define the composite function

$$v(M) \equiv -u(-G(M))$$

It is increasing and convex, and its first derivative is the utility of the consumption forgone when aggregate health care rises by one unit. The sets of untreated and treated types under private health care insurance are[2]

$$A_0(M) \equiv \{a \in A : h_1 - h_0 < v'(M)m\}$$

[2] The set of agents for whom $h_1 - h_0$ is just equal to $v'(M)m$ has measure zero, so the placement of these agents has no impact on any aggregate variable. Their placement in A_1 is an arbitrary choice of no consequence.

$$A_1(M) \equiv \{a \in A : h_1 - h_0 \geq v'(M)m\}$$

Then:

Lemma 1. *Assume that $G(\overline{M})$ is negative. The equilibrium values of C and M are unique. M is equal to zero and no insurance is sold if*

$$F(A_1(0)) = 0$$

M is positive and every agent purchases insurance if

$$F(A_1(0)) > 0$$

M is equal to $\overline{M}$ (i.e, every type is treated) if and only if

$$F(A_1(\overline{M})) = 1$$

4. The Public Health Care System

Now assume that all health care is provided by the government. It divides the types into two groups, those that will be treated and those that will not. It provides treatment by purchasing health care from the health care producers. Since the producers are competitive (and hence equate price to marginal cost), the total cost of health care is $G'(M)M$. The government finances its purchase of health care by imposing a lump-sum tax on each agent. The collective after-tax income of the agents is equal to C, which is just enough to enable then to purchase all of the available consumption goods. Each agent has an equal share of income, so (4) again holds.

The government designs the health care system to maximize an ex post social welfare function W that embodies some non-negative degree of inequality aversion:

$$W = \int_A U^\alpha dF = \int_{A_0} (h_0 + u(c))^\alpha dF + \int_{A_1} (h_1 + u(c))^\alpha dF \tag{5}$$

The degree of inequality aversion is measured by $\alpha \in (0, 1]$. There is no inequality aversion if α is equal to one, but as α falls toward zero, the degree of inequality aversion becomes progressively larger. The optimal health care system maximizes (5) subject to (2) and (4).

A necessary condition for the maximization of W is that the government treat only the agents for whom the benefit—the increase in the value of W—per unit of health care is greatest. Consequently, the sets of untreated and treated agents take the form[3]

$$A_0(k, M) \equiv \{a \in A : (h_1 - v(M))^{\alpha} - (h_0 - v(M))^{\alpha} < km\}$$
$$A_1(k, M) \equiv \{a \in A : (h_1 - v(M))^{\alpha} - (h_0 - v(M))^{\alpha} \geq km\} \tag{6}$$

Here, k is the government's policy instrument: the government treats only agents for whom the benefit per unit of health care is at least k.

The government's choice of k determines the remaining variables. Specifically, c and M are determined by conditions (2), (4) and (6). Since (4) shows that c is determined entirely by M, it is useful to focus on the relationship between k and M. Define the function

$$\Psi(M; k) \equiv \int_{A_1(k,M)} m dF$$

Aggregate health care M under any policy k is a fixed point of the equation

$$M = \Psi(M; k) \tag{7}$$

Lemma 2. *There is at least one solution to (7) for every non-negative k. If there is a unique solution for each k, M does not rise as k rises.*

There is some uncertainty about the uniqueness of the fixed point because an increase in M reduces consumption, expanding the set of types that qualify for treatment under any policy k and causing Ψ to rise. Thus, there could be several fixed point under any given k. A sufficient condition for uniqueness is that, for each k,

$$\frac{\partial \Psi}{\partial M} < 1 \qquad \text{for all } M \in [0, \overline{M}] \tag{8}$$

This condition can be interpreted as restricting the density of the agents on the sample space (so that a decrease in c does not shift many agents into the treatment group), or as placing a sufficiently tight upper bound on $G'(M)$ (so that an increase in M does not greatly reduce c). It will be assumed henceforth that a unique M is associated with each k.

Under this uniqueness assumption, M is a continuous non-increasing function of k, and c is a continuous non-decreasing function of k:

$$M = \widetilde{M}(k)$$

[3]The placement of marginal agents again has no impact on the aggregate variables C, M and W. Their assignment to the treatment group is arbitrary and without consequence.

$$c = \widetilde{c}(k)$$

The former function can be used to express the treatment groups in terms of k alone:

$$\widetilde{A}_i(k) \equiv A_i(k, \widetilde{M}(k)) \qquad i = 0, 1$$

The value of k determines the stringency of the government's treatment criteria. There exists some strictly positive $\underline{k}$ below which everyone is treated:

$$F\left(\widetilde{A}_1(k)\right) = 1 \qquad \text{for all } k \in [0, \underline{k}]$$

Likewise, there exists a finite $\overline{k}$ above which no-one is treated:[4]

$$F\left(\widetilde{A}_1(k)\right) = 0 \qquad \text{for all } k \geq \overline{k}$$

The connectedness of A implies that, between these values, $\widetilde{A}_1(k)$ contracts as k rises. It follows that $\widetilde{M}(k)$ is decreasing on the interval $(\underline{k}, \overline{k})$, and that $\widetilde{c}(k)$ is increasing over the same interval. Now consider the government's optimal policy.

Lemma 3. *Let k^* be the policy that maximizes social welfare, and let $\mu(k)$ be the marginal social value of consumption under the policy k:*

$$\begin{aligned} \mu(k) &\equiv \left.\frac{\partial W}{\partial c}\right|_k \\ &= \alpha u'(\widetilde{c}(k)) \left[\int_{\widetilde{A}_0(k)} (h_0 + u(\widetilde{c}(k)))^{\alpha-1}\, dF + \int_{\widetilde{A}_1(k)} (h_1 + u(\widetilde{c}(k)))^{\alpha-1}\, dF \right] \end{aligned}$$

If $\underline{k} < k^* < \overline{k}$, k^* satisfies the condition

$$k = \mu(k) G'(\widetilde{M}(k))$$

The social benefit of an additional unit of health care is k (by definition). The social cost of a unit of health care is $\mu G'(M)$, the social value of the consumption lost when one more unit of health care is provided. If the social benefit is greater than the social cost, social welfare can be increased by reducing k (i.e, relaxing the requirements for treatment), which increases aggregate health care. Likewise, social welfare rises with k if the social cost is greater than the social benefit. There is no small change in k that raises welfare when the social cost and social benefit of additional health care are equal.

[4] The number $\overline{k}$ is bounded above by the maximum value of $(h_1 - h_0)/m$, which is finite by the compactness of A.

4.1. Health Care without Inequality Aversion

If the social welfare function does not display inequality aversion ($\alpha = 1$), the treatment group under the optimal policy is simply

$$A_1 = \{a : h_1 - h_0 \geq v'(M)m\}$$

which is also the treatment group under private health care insurance. The allocation of resources under the optimal public insurance program is exactly the same as the allocation under private insurance. The only rationale for government involvement in health care in this model is inequality aversion. If there is not inequality aversion, it does not matter whether health care is a public or private institution.

Suppose that there is no aversion to inequality, and that there is a demand for private insurance. Would acceding to that demand be welfare improving? It would, but only because it would signal that the government's current health care is not optimally designed. Optimally restructuring that program would have the same impact as introducing private insurance.

4.2. Health Care with Inequality Aversion

Resources are allocated differently under public and private health care if there is inequality aversion. Figure 1 shows a cross section of the sample space (m is fixed at an arbitrary value) under the assumption that M is the same under both private and public care. Under private insurance, an agent is treated if his type places him above the EU-max locus. Under public insurance, an agent is treated if his type places him above the W-max locus. Since W-max cuts EU-max from below, there are types that are treated under private insurance but not under public insurance.[5] These types are contained in the darkly shaded region. Likewise, there are types that are treated under public insurance but not under private insurance.

Replacing private insurance with public insurance shifts health care away from those who have good health, whether they are treated or not, and towards those

[5] The EU-max locus has a slope of 1 while the slope of the W-max locus is

$$\frac{dh_1}{dh_0} = \left(\frac{h_1 - c}{h_0 - c}\right)^{1-\alpha} > 1$$

The loci do not necessary intersect in every cross section; but since total health care is assumed to be the same under each regime, they must intersect in some cross sections.

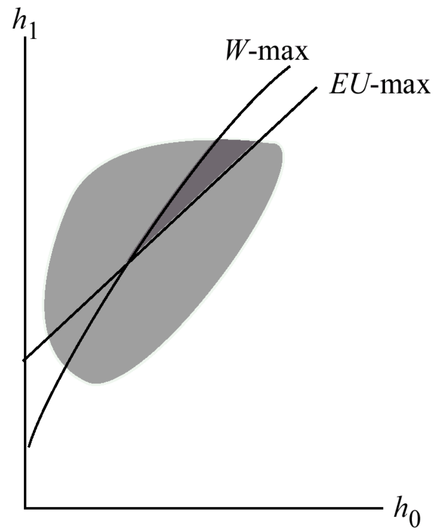

Figure 1: A cross section of the sample space. Types above EU-max would be treated under private insurance; types above W-max would be treated under public insurance.

who have poor health, whether treated or not. If there is inequality aversion, public insurance exists precisely to engineer this kind of reallocation of health care resources.

It is difficult to compare the resource allocations implied by these two systems, but the range of outcomes can be illustrated by two examples. The first example assumes that

$$h_1 = m = 1$$

and that, within the population, h_0 is uniformly distributed on the interval $[0, 1]$. Imagine that some but not all types are treated under private health care insurance. The agents purchase insurance that provides treatment if and only if h_0 is less than $1 - v'(M^\circ)$, where M° satisfies

$$M = \int_0^{1-v'(M)} dh_0$$

or

$$M = 1 - v'(M) \tag{9}$$

The solution to this equation is unique. Now consider public health care. An agent is treated if and only h_0 lies on some interval $[0,t]$. Social welfare is

$$\omega(t) = t[1 - v(M(t))]^\alpha + \int_t^1 [h_0 - v(M(t))]^\alpha dh_0$$

where

$$M(t) = t$$

Then

$$\begin{aligned}\omega'(t) &= [1 - v(M(t))]^\alpha - [t - v(M(t))]^\alpha - \\ &\quad \alpha v'(M(t)) \left\{ t[1 - v(M(t))]^{\alpha-1} + \int_t^1 [h_0 - v(M(t))]^{\alpha-1} dh_0 \right\}\end{aligned}$$

Let t° be the value of t under private health care insurance:

$$t^\circ \equiv 1 - v'(M^\circ) \tag{10}$$

Evaluating $\omega'(t)$ at t° gives, after some manipulation,

$$\omega'(t^\circ) = \alpha(1 - t^\circ)t^\circ \left\{ \left[\frac{1}{1 - t^\circ}\right] \int_{t^\circ}^1 [h_0 - v(M^\circ)]^{\alpha-1} dh_0 - [1 - v(M^\circ)]^{\alpha-1} \right\} > 0$$

It follows that the optimal public health care policy sets t above t°, so that aggregate health care is greater under the public health care system than under private health insurance.

The difference between these two systems is, of course, the way in which the treatment costs are allocated. Equation (10) states that an agent of type $(t^\circ, 1, 1)$ feels that the benefit of treatment is just offset by the cost of treatment. A public health system spreads this cost across all of society. Since an agent of this type has a lower utility than any other agent,[6] spreading the cost of his treatment across the whole of society raises social welfare. Thus, social welfare rises when the agent is treated under public health care, but does not change when he is treated under private health insurance.

The second example assumes that

$$h_0 = 0, m = 1$$

[6] An agent with a smaller h_0 is treated and therefore healthier; an agent with a larger h_0 is healthier even though he is also untreated.

and that h_1 is uniformly distributed on $[0,1]$. Under private health care insurance, every agent purchases insurance that provides treatment if and only if h_1 is greater than $v'(M^\circ)$, where M° is again the solution to (9).[7] Under public health care, an agent is treated if and only if h_1 lies within some interval $[t,1]$. Social welfare is now

$$\omega(t) = -tv(M(t))^\alpha + \int_t^1 [h_1 - v(M(t))]^\alpha dh_1$$

where

$$M(t) = 1 - t$$

Then

$$\omega'(t) = -v(M(t))^\alpha - [t - v(M(t))]^\alpha + \alpha v'(M(t)) \left\{ \int_t^1 [h_1 - v(M(t))]^{\alpha-1} dh_1 - tv(M(t))^{\alpha-1} \right\}$$

As before, let t° be the value of t under private health care insurance:

$$t^\circ \equiv v'(M^\circ)$$

Evaluating the derivative at t° gives

$$\omega'(t^\circ) = \alpha t^\circ \left\{ \int_{t^\circ}^1 [h_1 - v(M(t))]^{\alpha-1} dh_1 - t^\circ v(M^\circ) - \left(\frac{1}{t^\circ}\right) \int_0^{t^\circ} [h_1 - v(M(t))]^{\alpha-1} dh_1 \right\}$$

Inspection of this expression shows that the sign of the derivative is not certain. It is negative if t° is near zero and it is positive if t° is near one, so that the optimal public health system plan might treat either more types or fewer types than the private health care system. Suppose, for example, that the utility function is

$$u(c) = c$$

and that the production possibility frontier is

$$c + \beta M = 1$$

[7] As before, the fraction of the types that are treated is $1 - v'(M^\circ)$.

t°	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.86
t^*	0	0.087	0.176	0.271	0.374	0.488	0.616	0.757	0.907	1.0

Table 1: The coverage of private health care insurance and the optimal public health care system in the second example

Then t° is equal to β, while the socially optimal value of t (call it t^*) depends upon both α and β. Table 1 shows the relationship between t^* and t° when α is equal to 0.6. The public health care plan is more extensive than the private health care plan (as it was in the first example) if β is smaller than some value near 0.55; but if β is larger than this value, the private health care plan is less extensive.

As in the first example, the relationship between t° and t^* depends upon the welfare effect of spreading one agent's treatment cost across all of society. If t° is small, an agent of type $(1, t^\circ, 1)$ would be less healthy—and have lower utility—than most members of society, even after he has been treated.[8] Shifting his treatment cost to the whole of society raises social welfare. Consequently, treating agents of this type under the public system raises social welfare, even though treating them under private insurance has no impact on welfare. The conclusion is reversed if t° is large. An agent of type $(1, t^\circ, 1)$ would then be healthier than most members of society after he has been treated. Shifting his costs of treatment to the whole of society reduces social welfare. It follows that treating an agent of this type under the public system reduces social welfare, while treating him under private insurance has no impact on social welfare.

It seems that anything goes. The optimal public health care system might treat everyone who would be treated under private health care, or it might not. Aggregate health care might be greater under the optimal health care system than under private health care, or it might not.

5. A System with Both Private and Public Insurance

Assume that there is aversion to inequality ($\alpha < 1$), and imagine that a public health system is in place and that private health care insurance is allowed. The private insurers offer policies that cover treatment for types that are not covered

[8] He would be healthier than every untreated agent but less healthy than every treated agent. Since t° is small, the second group of agents outnumbers the first group.

under the public system. The government must recognize the response of the private insurers when it designs its own system.

The government finances its program by imposing a lump-sum tax on each agent. If the private insurers are viable (they need not be), each agent will also choose to purchase private insurance, and each agent will pay the same premium. Consequently, each agent's consumption is again determined by (4).

The set A is now split into three subsets: the set of untreated agents A_0, the set of agents treated under public insurance A_1, and the set of agents treated under private insurance A_2. As before, the government maximizes social welfare by treating the agents for whom the increase in social welfare per unit of health care is greatest, so that the set of types treated under its program is again given by (6). The private insurers will cover all of the remaining types for which the benefit of treatment exceeds the cost:

$$A_2(k, M) \equiv \{a \in A \backslash A_1(k, M) : h_1 - h_0 \geq v'(M)m\} \tag{11}$$

All other types are untreated.

Aggregate health care is the sum of the health care provided by the government and the health care provided by private insurers:

$$M = \int_{A_1(k,M)} m dF + \int_{A_2(k,M)} m dF \tag{12}$$

For each k, an equilibrium consists of a solution to (6), (11) and (12).

Define the function

$$\Gamma(M; k) \equiv \int_{A_2(k,M)} m dF$$

An increase in M either expands $A_1(k, M)$ or leaves it unchanged, and an increase in M either increases $G'(M)$ or leaves it unchanged. An expansion of A_1 or an increase in $v'(M)$ tends to shrink $A_2(k, M)$, so $A_2(k, M)$ either contracts of remains unchanged as M rises. Thus $\Gamma(M; k)$ is non-increasing in M.

Aggregate health care M under any policy k is a fixed point of the equation

$$M = \Psi(M; k) + \Gamma(M; k) \tag{13}$$

A fixed point exists, and (8) is a sufficient condition for its uniqueness.

Lemma 4. *If the fixed point of (13) is unique for each k, M is a non-increasing function of k. Let this function be*

$$M = \widehat{M}(k)$$

The set

$$\widehat{A}_1(k) \equiv A_1(k, \widehat{M}(k))$$

either contracts or remains unchanged as k rises, and the set

$$\widehat{A}_2(k) \equiv A_2(k, \widehat{M}(k))$$

either expands or remains unchanged as k rises.

Less can be said about the set of all treated types, $\widehat{A}_1(k) \cup \widehat{A}_2(k)$. An increase in k transfers some types from A_1 to A_2; but some of the types added to A_2 were not previously included in A_1 and some of the types dropped from A_1 are not added to A_2. Even though aggregate health care $\widehat{M}$ is non-increasing in k, it is not certain whether the measure of all treated types rises or falls as k rises.

Lemma 5. *Assume that the fixed point of (13) is unique for each k, and let M'' be aggregate health care under a private health care system. Then there exists some $k' \geq \underline{k}$ such that $\widehat{A}_2(k)$ is empty for all $k \leq k'$, and some $k'' > \overline{k}$ such that $\widehat{A}_1(k)$ is empty for all $k \geq k''$. Also,*

1. *$\widehat{M}(k) = \widetilde{M}(k)$ for all $k \leq k'$ and $\widehat{M}(k) = M''$ for all $k \geq k''$.*
2. *$\widehat{M}(k)$ is decreasing on the interval (k', k''), and $\widehat{M}(k) > \widetilde{M}(k)$ on the same interval.*

Figure 2 illustrates these results. The public health care system treats all types when $k \leq \underline{k}$. The coverage of this system contracts as k rises above $\underline{k}$ until, when k reaches k', private insurers begin to provide supplementary insurance. Further increases in k result in greater private coverage and reduced public coverage. The public program disappears at k'', and there is only private coverage at all greater values of k. At every k greater than k', aggregate health care is greater under parallel systems of health care than under public health care alone.

An optimally designed public health care system is characterized by a value of k that lies somewhere between $\underline{k}$ and $\overline{k}$. Can this value be so low that there would be no latent demand for private health care? Can it be so high that aggregate health care is less than it would be under private health care? Yes, it can, as the examples of section 4.2 have shown.

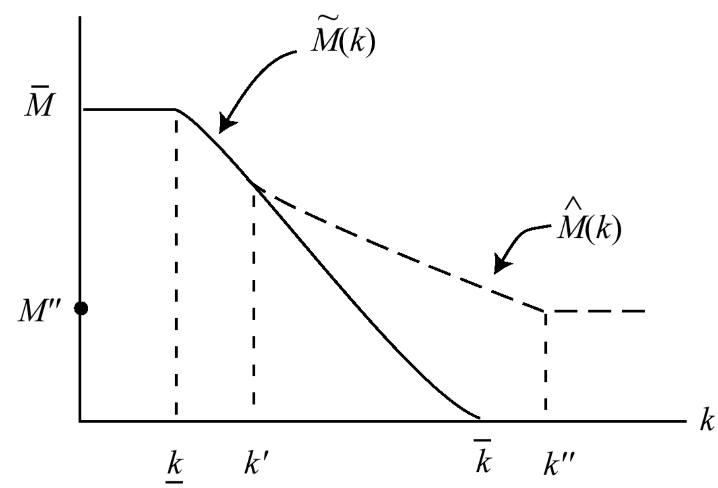

Figure 2: Aggregate health care under a public health care system, and under a combined public and private health care system.

5.1. Welfare Implications

A parallel system of private health insurance reduces welfare:

> Assume that both private and public health care are permitted, and that the private insurers offer some or all of the coverage ($k \geq k'$). Then there is a public health care system under which social welfare W is higher.

Suppose that both private and public health care are permitted, and that the government sets k at some k^* that is greater than k'. Aggregate health care is then equal to $\widehat{M}(k^*)$. If private health care is not permitted, there is a smaller value of k (call it k^{**}) such that $\widetilde{M}(k^{**})$ is equal to $\widehat{M}(k^*)$. The only difference between the "mixed" equilibrium at k^* and the purely public equilibrium at k^{**} is the manner in which health care is allocated. All of the health care is allocated to maximize W in the purely public equilibrium, but only a part of the health care is allocated to this end in the mixed equilibrium. It follows that maximal social welfare is higher when private health care insurance is not permitted.

This result suggests that the role of private insurance in an inequality-averse society is extremely limited.

- If the government chooses the optimal value of k, and if that k is no greater than k', there is no demand for private health care insurance.

- If the government chooses the optimal value of k, and if that k is above k', there is a demand for private health care insurance. Acceding to that demand would necessarily reduce welfare. Aggregate health care would be greater if private insurance were permitted (and k were not changed), but it would be both too large and badly allocated.

- If the government sets k at a value that is not optimal, there might (or might not) be a demand for private health care insurance. Acceding to this demand would always be a worse strategy than simply optimizing the government program.

If the social welfare function displays an aversion to inequality, the observation that there is a demand for private insurance does not imply that its introduction would be welfare improving. If the public program is optimally designed, introducing private insurance would lower welfare; if the pubic program is not optimally designed, improving the public program would yield a larger increase in welfare.

5.2. Second-Best Policy

How would the public health care program adjust to the introduction of private health care insurance? This question has usually been answered by imagining that there is an innate difference in efficiency between the public and private sectors, or by investigating the manner in which the incentives of doctors and other health care providers are changed. This section offers an alternative approach. The private sector allocates health care in a manner that does not maximize social welfare. The public sector might be able to influence the extent of the misallocation by strategically altering its own program. This possibility is described by the following lemma.

Lemma 6. *Assume that both public and private health care is permitted. Let $\mu(k)$ be the marginal social value of consumption:*

$$\mu(k) = \alpha u'(\widehat{c}(k)) \left[\int_{\widehat{A}_0(k)} (h_0 - u(\widehat{c}(k)))^{\alpha-1} dF + \int_{\widehat{A}_1(k) \cup \widehat{A}_2(k)} (h_1 - u(\widehat{c}(k)))^{\alpha-1} dF \right]$$

where

$$\widehat{c} \equiv -G(\widehat{M})$$

Let k^* be the welfare-maximizing value of the instrument k, and assume that both the private and public insurers are active at k^* (that is, $k' < k^* < k''$). Then

$$k^* = \mu(k^*)G'(\widehat{M}(k^*))$$

if G'' is equal to zero , and

$$k^* < \mu(k^*)G'(\widehat{M}(k^*))$$

if G'' is positive.

If the marginal cost of health care is constant at the optimum, the rule that characterizes the optimal design of the public program is the same whether there are private insurers or not. Specifically, the social value of another unit of health care, k, is equal to the social value of the consumption that must be given up to obtain that unit of health care, $\mu G'(M)$.[9] However, this rule does not characterize the optimal design of the public program if the marginal cost of health care is increasing and the private insurers are active.

The reasoning behind this result can be understood by referring once again to the cross section in Figure 1. If G'' is zero in the neighbourhood of k^*, a small change in k leaves the position of the EU-max locus unchanged. Every type that lies above the locus is treated under one of the two programs. The only role of the instrument k is to divide the types that lie below this locus into two groups, those who are treated under the public program and those who are untreated. A decrease in k shifts W-max downward, shifting some types from the untreated group to the public treatment group. The social value of additional health care is equal to its social opportunity cost under the optimal policy. Now suppose that G'' is positive in the neighbourhood of k^*. A decrease in k shifts W-max downward; but it also shifts EU-max upward, so that some types are moved from the private treatment group to the untreated group. From the perspective of the public insurer (which conscientiously maximizes W), the latter shift is welfare improving: the private insurers allocate resources badly, so social welfare rises when their program is curtailed. The optimally designed public program exploits this effect by pushing k below $\mu G'(M)$.

[9] The form and interpretation of the rule are the same whether or not private insurers are allowed. However, the optimal values of k and M differ in these two cases because the functions that relate μ and M to k differ.

6. Conclusions

An economy in which the public health system is designed to mitigate inequality has been examined. Replacing private health care insurance with this kind of public system shifts health care away from those who are relatively health with or without treatment and towards those who are relatively unhealthy with or without treatment. Aggregate health care can either rise or fall. If such a system is in place, introducing a parallel system of private insurance cannot raise welfare. If the insurers are active, their presence must reduce welfare.

An important limitation of these results is that the agents were assumed to be ex ante identical, and in particular, to have equal claims on output. It is not clear what impact the introduction of ex ante income disparities would have on the results. These disparities would create greater inequality across the population, so that the welfare gain to using health care as a redistributive mechanism rises. At the same time, income disparities lead to different demands for health care, so that a "one size fits all" public health care system entails welfare losses. The net effect on welfare is difficult to predict. However, if there are income disparities, it might be misleading to draw inferences from a model in which health care is the only redistributive instrument. One could imagine using the tax system to reduce the inequalities that arise from ex ante income disparities, and the public health care system to reduce the disparities that arise from ex post differences in health status. The practice of combining ex ante and ex post redistributive instruments is already familiar to us; for example, a large amount of redistribution occurs through progressive income taxation, but there is further redistribution towards those who find themselves unemployed or disabled.

A. Appendix

Proof of Lemma 1: Define the function

$$\Phi(M) \equiv \int_{A_1(M)} m dF$$

An increase in M either contracts $A_1(M)$ or leaves it unchanged, so Φ is non-increasing in M. By (2) the equilibrium value of M is the fixed point of

$$M = \Phi(M)$$

Since Φ is non-increasing, the fixed point is unique. If $F(A_1(0)) = 0$, $\Phi(0) = 0$ so that the equilibrium values of M and C are 0 and $-G(0)$. If $F(A_1(0)) > 0$,

$\Phi(0) > 0$, so the equilibrium value of M must be positive. There are then two cases. First, if $F(A_1(\overline{M})) < 1$, $\Phi(\overline{M}) < \overline{M}$, implying that the fixed point must satisfy the condition $0 < M < \overline{M}$. Second, if $F(A_1(\overline{M})) = 1$, $F(A_1(M)) = 1$ for all $M \in [0, \overline{M}]$ and hence $\Phi(M) = \overline{M}$ for all $M \in [0, \overline{M}]$. The fixed point is then $\overline{M}$. In both of these cases, the equilibrium value of C is $-G(M) > 0$. ■

Proof of Lemma 2: An increase in k leaves $A_1(k, M)$ unchanged or shrinks it, so Ψ is non-increasing in k. An increase in M leaves $A_1(k, M)$ unchanged or expands it, so Ψ is non-decreasing in M. For each k, M is the fixed point of

$$M = \Psi(M; k)$$

$\Psi(M)$ is continuous and maps $[0, \overline{M}]$ into $[0, \overline{M}]$, so at least one solution exists. Since Ψ is non-decreasing in M, multiple solutions are possible. If the solution is unique, an increase in k either leaves the graph of Ψ (against M) unchanged or shifts it downwards, so that M is unchanged or lower. ■

Proof of Lemma 3: Social welfare under any k° can be expressed as the difference between social welfare when the treatment group is fixed at $\widetilde{A}_1(k)$, and the social welfare lost because the actual treatment group $\widetilde{A}_1(k^\circ)$ differs from $\widetilde{A}_1(k)$.

$$W(k^\circ; k) = \sum_{i=0}^{1} \left(\int_{\widetilde{A}_i(k)} \left[h_i - v(\widetilde{M}(k^\circ)) \right]^\alpha dF \right) - J(k^\circ; k)$$

Here, $J(k^\circ; k)$ is the change in social welfare caused by the removal of agents from the treatment group as k° rises above the value k. $J(k^\circ)$ is positive when k° is greater than k, zero when k° is equal to k, and negative when k° is smaller than k. Setting the policy instrument equal to k maximizes social welfare if any deviation of k° from k reduces welfare, or equivalently, if $W(k^\circ; k)$ has a stationary point at k:

$$\left. \frac{\partial W(k^\circ; k)}{\partial k^\circ} \right|_{(k,k)} = 0$$

Differentiating W with respect to k° and evaluating the resulting expression at k gives

$$\left. \frac{\partial W(k^\circ; k)}{\partial k^\circ} \right|_{(k,k)} = -\mu(k) G'(\widetilde{M}(k)) \widetilde{M}'(k) - J'(k)$$

An arbitrarily small increase in k removes from the treatment group the agents who are marginal candidates for treatment under the current policy. The characteristic of these agents is that

$$[h_1 - v(\widetilde{M}(k))]^\alpha - [h_0 - v(\widetilde{M}(k))]^\alpha = km$$

Since the left-hand side of this equation is the social benefit of moving an agent into the treatment group, integrating over all of the agents moved *out* of the treatment group by an arbitrarily small increase in k gives

$$J'(k) = -k\widetilde{M}'(k)$$

Then

$$\left.\frac{\partial W(k^\circ;k)}{\partial k^\circ}\right|_{(k,k)} = \left[k - \mu(k)G'(\widetilde{M}(k))\right]\widetilde{M}'(k)$$

Since $\widetilde{M}$ does not switch signs, a stationary point only occurs when the bracketed expression is equal to zero. ■

Proof of Lemma 4: At each M, an increase in k tends to contract $A_1(k,M)$. The types removed from $A_1(k,M)$ are not always added to $A_2(k,M)$, so $A_1 \cup A_2$ tends to contract as k rises, causing the graph of $\Psi + \Gamma$ to shift downward. Consequently, the fixed point $\widehat{M}(k)$ either falls or remains unchanged. Now consider $A_1(k,\widehat{M}(k))$. An increase in k tightens the criteria for treatment under the public program, so that $A_1(M,k)$ tends to contract. This effect is reinforced by a fall in $\widehat{M}(k)$. Any contraction of $A_1(k,\widehat{M}(k))$ leaves more scope for private insurers, so that $A_2(k,\widehat{M}(k))$ tends to rise. There is an additional expansion of $A_2(k,\widehat{M}(k))$ if $G'(\widehat{M}(k))$ falls as k rises. ■

Proof of Lemma 5: The existence of k' follows from the observations that $\widehat{A}_2(k)$ does not expand as k falls, and that $\widehat{A}_2(\underline{k}) = \varnothing$ (because $\widehat{A}_1(\underline{k}) = A$). Likewise, the existence of k'' follows from the observation that a sufficiently tight standard for public treatment excludes every type. The inequality $k'' > \overline{k}$ follows from the observation that k'' is the lowest k at which

$$F\left(A_1(k,M'')\right) = 0$$

while $\overline{k}$ is the lowest k at which

$$F\left(A_1(k,0)\right) = 0$$

There is only public health care at every k below k', so $\widehat{M}(k) = \widetilde{M}(k)$. There is only private health care at every k above k'', so $\widehat{M}(k) = M''$. The connectedness of A implies that $\widehat{M}(k)$ is decreasing over the intermediate range. Every increase in k shifts the graph of $\Psi+\Gamma$ downward, causing the value of the fixed point $\widehat{M}(k)$ to decline. ■

Proof of Lemma 6: Let $\widehat{M_1}(k)$ and $\widehat{M_2}(k)$ be the aggregate quantities of health care provided by the public and private insurers respectively. Social welfare under any policy k° can be written as

$$W(k^\circ;k) = \int_{\widehat{A}_0(k)} \left[h_0 - v(\widetilde{M}(k^\circ))\right]^\alpha dF + \int_{\widehat{A}_1(k)\cup\widehat{A}_2(k)} \left[h_{_1} - v(\widetilde{M}(k^\circ))\right]^\alpha dF - J(k^\circ;k) + L(k^\circ;k)$$

The first two terms describe social welfare when the public and private treatment groups are artificially fixed at $\widehat{A}_1(k)$ and $\widehat{A}_2(k)$ respectively. The next term describes the change in welfare induced by changing the public treatment group from $\widehat{A}_1(k)$ to $\widehat{A}_1(k^\circ)$, and the last term describes the change in welfare induced by changing the private treatment group from $\widehat{A}_2(k)$ to $\widehat{A}_2(k^\circ)$. Since $\widehat{A}_1(k^\circ)$ shrinks as k° rises, $J(k^\circ;k)$ is positive (negative) when k° is greater than (less than) k. Since $\widehat{A}_2(k^\circ)$ expands as k° rises, $L(k^\circ;k)$ is also positive (negative) when k° is greater than (less than) k. (Note that these two terms enter $W(k^\circ;k)$ with opposite signs.) Setting the policy instrument equal to k maximizes social welfare if any deviation of k° from k reduces welfare, which requires $W(k^\circ;k)$ to have a stationary point at k:

$$\left.\frac{\partial W(k^\circ;k)}{\partial k^\circ}\right|_{(k,k)} = 0$$

Differentiating W with respect to k° and evaluating the resulting expression at k gives

$$\left.\frac{\partial W(k^\circ;k)}{\partial k^\circ}\right|_{(k,k)} = -\mu(k)G'(\widehat{M}(k))\widehat{M}'(k) - J'(k) + L'(k)$$

An arbitrarily small increase in k removes from $\widehat{A}_1(k)$ the agents who are marginal candidates for treatment under the policy k. The characteristic of these agents is that

$$[h_1 - v(\widehat{M}(k))]^\alpha - [h_0 - v(\widehat{M}(k))]^\alpha = km$$

Since the left-hand side of this equation is the social benefit of moving an agent into the treatment group, integrating over all of the agents moved *out* of the treatment group by an arbitrarily small increase in k gives

$$J'(k) = -k\widehat{M}_1'(k)$$

Then

$$\left.\frac{\partial W(k^\circ;k)}{\partial k^\circ}\right|_{(k,k)} = \left[k - \mu(k)G'(\widehat{M}(k))\right]\widehat{M}_1'(k) + \left[L'(k) - \mu(k)G'(\widehat{M}(k))\widehat{M}_2'(k)\right]$$

or

$$\left.\frac{\partial W(k^\circ;k)}{\partial k^\circ}\right|_{(k,k)} = \left[k - \mu(k)G'(\widehat{M}(k))\right]\widehat{M}'(k) + \left[L'(k) - k\widehat{M}_2'(k)\right] \tag{14}$$

Consider the two cases in turn.

Suppose that $G''(k) = 0$. Inspection of (11) shows that an increase in k causes $\widehat{A}_2(k)$ to expand only because some of the types removed from $\widehat{A}_1(k)$ are added to $\widehat{A}_2(k)$. It follows that all of the agents added to $\widehat{A}_2(k)$ have the same characteristic as the agents deleted from $\widehat{A}_1(k)$:

$$[h_1 - v(\widehat{M}(k))]^\alpha - [h_0 - v(\widehat{M}(k))]^\alpha = km \tag{15}$$

Integrating over all of the agents moved into $\widehat{A}_2(k)$ by an arbitrarily small increase in k gives

$$L'(k) = k\widehat{M}_2'(k)$$

Substituting this expression into (14) shows that the stationary point occurs where

$$k = \mu(k)G'(\widehat{M}(k))$$

Now suppose that $G'' > 0$. When k rises by some arbitrarily small amount, some types are moved from $\widehat{A}_1(k)$ to $\widehat{A}_2(k)$, as before, but others are transferred from $\widehat{A}_0(k)$ to $\widehat{A}_2(k)$. (The latter transfer occurs because an increase in k reduces aggregate health care $\widehat{M}(k)$, which in turn reduces the marginal cost of health care $G'(\widehat{M}(k))$. Inspection of (11) shows that, for any given $\widehat{A}_1(k)$, the reduction in the marginal cost of health shifts some types into $\widehat{A}_2(k)$.) The types at the boundary between private and public health care satisfy (15) while the types at the boundary between private health care and no health care satisfy the condition

$$h_1 - h_0 = km$$

For the latter types,

$$[h_1 - v(\widehat{M}(k))]^\alpha - [h_0 - v(\widehat{M}(k))]^\alpha < (h_1)^\alpha - (h_0)^\alpha < h_1 - h_0 = km$$

Since $L'(k)$ is the increase in social welfare generated by treating all of the agents added to $\widehat{A}_2(k)$ when k rises marginally,

$$L'(k) < k\widehat{M_2}(k)$$

Inspection of (14) shows that the condition

$$k < \mu(k)G'(\widehat{M}(k))$$

characterizes the stationary point. ■

References

[1] Boadway, R., M. Leite-Monteiro, M. Marchand and P. Pestieau (2006), Social insurance and redistribution with moral hazard and adverse selection, Scandinavian Journal of Economics 108:2, 279–298.

[2] Boadway, R., M. Leite-Monteiro, M. Marchand, and P. Pestieau, P. (2003), Social insurance and redistribution. In S. Cnossen and H.-W. Sinn (eds), Public Finance and Public Policy in the New Century, 333–358. Cambridge, MA: MIT Press, Cambridge, MA.

[3] Cremer, H. and P. Pestieau (1996), Redistributive taxation and social insurance, International Tax and Public Finance 3, 281–295.

[4] Hammond, P. J. (1981), Ex-ante and ex-post welfare optimality under uncertainty, Economica, New Series 48:191, 235–250.

[5] Harris, R. G. (1978), Ex-post efficiency and resource allocation under uncertainty, Review of Economic Studies 45:3, 427–436.

[6] Mirrlees, J. A. (1974), Notes on welfare economics, information and uncertainty. In M. Balch, D. McFadden and S. Wu (eds), Essays on Economic Behaviour under Uncertainty, 243–258. Amsterdam: North Holland.

[7] Pauly, M. V. (1971), Medical Care at Public Expense: A Study in Applied Welfare Economics. New York: Praeger.

[8] Petretto, A. (1999), Optimal social health insurance with supplementary private insurance, Journal of Health Economics 18, 727–745.

[9] Rochet, J.-C. (1991), Incentives, redistribution and social insurance, Geneva Papers of Risk and Insurance 16, 143–165.

[10] Starr, R. M. (1973), Optimal production and allocation under uncertainty, Quarterly Journal of Economics 87:1, 81–95.

[11] Tobin, J. (1970), On limiting the domain of inequality, Journal of Law and Economics 13: 2, 263–277.